DETERMINATION OF SAMPLE SIZE

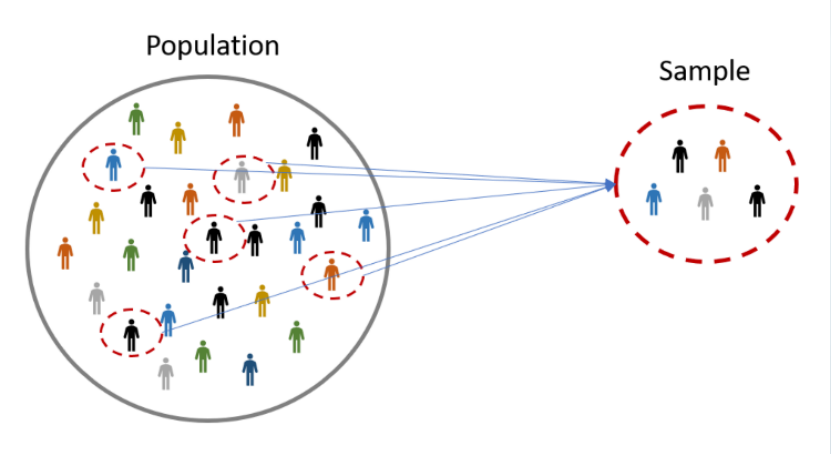
by Aazifah Ilham[1], Nurul Iman Hafizah[1], Boon-How Chew[1,2]

[1]Clinical Research Unit, Hospital Pengajar Universiti Putra Malaysia (HPUPM Teaching Hospital), Persiaran Mardi-UPM, 43400 Serdang, Malaysia
[2]Department of Family Medicine, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Selangor, Malaysia.

One of the most crucial aspects in conducting clinical research is the determination of sample size. This is because adequate sample size for any newly proposed study is based on the research question so that it has to be powered to a sufficient precision level at least to answer the primary objective of the proposed study. This has always been one of the fundamental requirements by clinical ethic committees.

"Sample" refers to a set of individuals included in research which adequately represents the population from where it is drawn so that true inferences from the results can be made (see figure below). Quality and strength of statistical inference depends largely on the sample size selected. Health personnel need to be aware of the fact that the planning of any particular study includes the decision on how large a sample should be selected from a population. Estimation of sample size require in research is one of the most common reason for visits to statistician. There is no "one-size-fits-all" formula to determine the right sample size. Therefore, confusion is easily occurred due to many formulas with 'unknown' components in the calculation.

The aim of this article is to introduce readers the important components that they need to know before calculating sample size.



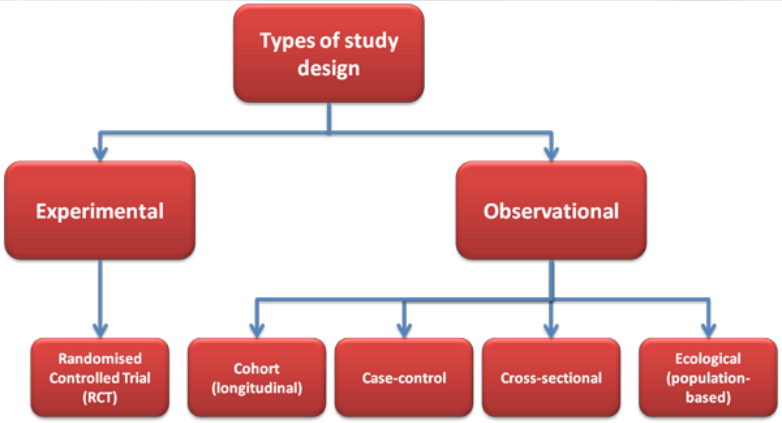
What we must know before calculating the sample size?

Generally, the sample size for any study depends on:

- The objectives and design of the study
- Types of variables
- Hypothesis testing
- Types of statistical analyses used
- Type of errors
- Power of the study
- Effect size
- Standard deviation in the population
- Margin of error

Types of study design

The objectives of the study are critical to determine the type of study design that is going to be conducted. It is important to note that different study de sign requires different method or formula to calculate the sample size. Types of study design is divided into experimental and non-experimental (observational) study design. Observational study design does not involve overt manipulation or management of variables. Example of observational study design are cross sectional, cohort, case control study design. Experimental involved in the manipulation and management of variable.



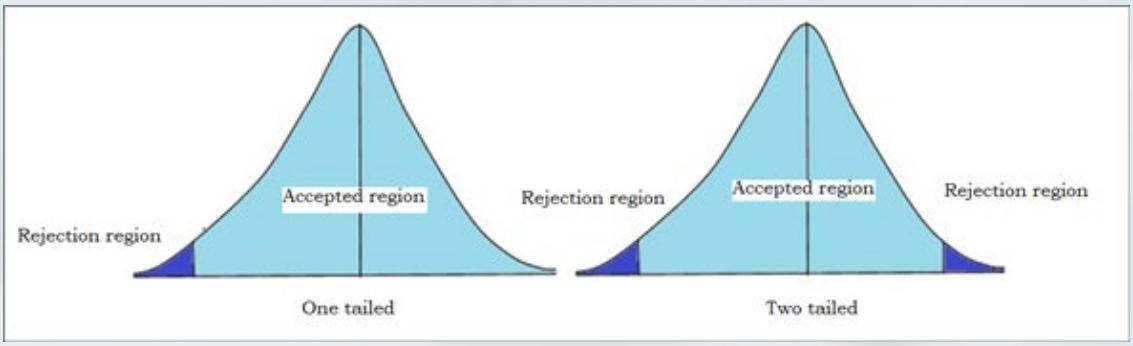
Types of variables

Variable can be divided into two groups which are numerical and categorical variables. Numerical variable can be divided into continuous and discrete. While categorical variable can be further divided into nominal and ordinal.

Another type of variable is dependent and independent variable. Independent variable is the variable that influence the outcome (dependent variable). It is necessary to decide on independent and dependent variable as well as the categories of variables are important as they influence the statistical test and sample size calculation.

Hypothesis testing

Most clinical research requires hypothesis testing. It is a method for testing a claim or hypothesis about parameter in a population, using data measured in a sample. There are two main types of hypotheses testing: one tail hypothesis (directional) and two tails hypotheses (non- directional). Two-tailed hypothesis is used when the alternative hypothesis is stated not equal while one-tailed test is used when the alternative hypothesis is stated greater or lesser a value stated in null hypothesis. Two-tailed test are generally employed because of statistical uncertainty that the result can go either direction. One-tailed test appears in non-inferiority trials or superiority trials where the outcome is obvious to be at one direction of effect by the nature of the intervention/treatment. However, in some cases, there is no hypothesis as the investigator just wants to observe the data descriptively. For example, in prevalent study where there is no hypothesis to be tested.



Statistical test chosen

Knowledge on statistical test is greatly needed prior to calculating sample size. Statistical test is used in hypothesis testing. They can be used to determine whether the predictor has statistical significance on an outcome variable. To produce ethically and scientifically valid result, your sample size needs to be statistically sufficient to approximate the true distribution of the population being studied. The sample size needs to fulfil the statistical assumption in order to make the statistical test valid. It is important to decide *a priori* which statistical test to use in the study before embarking in sampling calculation. For example, observational studies with large population size that involve logistic regression in the analysis, taking a minimum sample size of 500 is necessary to derive the statistics that represent the parameters [3]. Another recommended rule of thumb for logistic regression is the formula n = 100 + 50i where i refers to number of independent variables in the final model [3].

Type of errors

- Type I error (alpha) – The probability of falsely rejecting null hypothesis and detecting a statistically difference when the groups in reality are not different. The alpha most commonly fixed at 0.05 which means that the researcher desires a <5% chance of drawing a false positive. Reducing alpha will increase the sample size required.
- Type II error (beta) – The probability of falsely accepting null hypothesis and detecting a statistically significant when the groups in reality is different. Conventionally the beta is set at the level of 0.20 meaning that the researcher desires a <20% chance of false negative conclusion.

Power (1 – beta)

The "power" of the study is equal to (1 – β) which is the probability of failing to detect a difference when actually there is a difference. The power of a study increases as the chances of committing a Type II error decrease. The probability of correctly rejecting null hypothesis and detecting a statistically significant result when a specific difference between the group in reality exist. The power is the compliment of beta. So, if the beta is 0.20 the power is 0.80 or 80% representing the prob ability of avoiding a false negative conclusion or chance of correctly rejecting a null hypothesis. Standard normal variate for 80% power is 0.84 and for 90% is 1.28. Standard normal variate is used to transform the data to fit normal distribution. Usually, most clinical research accepts a power at least 80%. Reducing power will reduce the sample size required.

The smallest effect of interest/ Minimal clinically relevant difference/Effect size

The minimal difference between the groups is what the investigator considers biologically plausible and clinically relevant difference in the treatment effect between the comparing groups. They are the magnitude and direction of the different between two groups or the relationship between two variables. They can be a different between two means, percentages or correlations. Commonly used are raw group different, standardized mean difference and odd ratio. In the context of effect size, the sample size is inversely proportional in the required sample size, as such when effect size is large, a smaller sample size is needed and vice versa.

The variability

Researcher must foresee the population variance of a given outcome variable which usually express in the form of standard deviation (SD) in continuous variable. A larger variability in the variable of interest (of the primary objective of the proposed study), a larger sample size will be required to achieve the same power, and vice versa.

Margin of error

Margin of error is defined as half of the width of confident interval. It is usually needed in calculation of study not involved in hypothesis testing. A larger margin of error, if acceptable, in the variable of interest (of the primary objective of the proposed study), a smaller sample size is required, and vice versa.

Conclusion

There are a few methods to determine sample size which is either by using formula or dedicated software and websites (eg: GPower, OpenEpi etc). However, understanding in each component of the formulae is important to prevent error in computing the numbers as using any formula or software they can still give us any value that might be incorrectly computed.

Reference:

1/ Joseph A. Durlak, How to Select, Calculate, and Interpret Effect Sizes, *Journal of Pediatric Psychology*, Volume 34, Issue 9, October 2009, Pages 917–928, https://doi.org/10.1093/jpepsy/jsp004

2/ Noordzij M, Dekker FW, Zoccali C, Jager KJ. Sample size calculations. Nephron Clin Pract. 2011;118(4):c319-23. doi: 10.1159/000322830. Epub 2011 Feb 3. PMID: 21293154.

3/ Bujang MA, Sa'at N, Sidik TMITAB, Joo LC. Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy Between Statistics and Parameters Based on Real Life Clinical Data. Malays J Med Sci. 2018 Jul;25(4):122-130. doi: 10.21315/mjms2018.25.4.12. Epub 2018 Aug 30. PMID: 30914854; PMCID: PMC6422534.

4/ Majid U. Research fundamentals: Study design, population, and sample size. URNCST Journal. 2018 Jan 10: 2(1). https://urncst.com/index.php/urncst/article/view/16 DOI Link: https://doi.org/10.26685/urncst.16

5/Naing NN. Determination of sample size. The Malaysian Journal of Medical Sciences : MJMS. 2003 Jul;10(2):84-86. PMID: 23386802; PMCID: PMC3561892